

THE COMPUTER U F O NEWSLETTER

Volume 01 - Number 06

Computer UFO Newsletter

Vol. 1 Issue 6

The Computer UFO Newsletter" is an aperiodical independent newsletter published irregularly six times per volume, entirely devoted to the presentation of works and discussions about the use and applications of computer in ufology.

CUFON can be received by exchange with other UFO magazines or supplying a minimum fee of 15,000 Italian lire for six issues (surface mail: for air mail, add 4,000 lire). This subscription rate allows to receive all issues published in volume number one. Please send only international Postal Money Order payable to "Maurizio Verga": all checks are refused due to very high exchange costs.

This publication is entirely composed by suitable software on a Commodore 128 personal computer and printed by a MPS-802 printer. Word processor files relating to published articles are available on request.

Our general ADDRESS:

Computer UFO Newsletter

Maurizio Verga
via Matteotti 85
22072 Cermenate (Como)
ITALY

Views expressed by contributors are not necessarily shared by the Editor. Articles published in the Newsletter can be reprinted only with written permission by the Editor and giving full credit to the author and to "The Computer UFO Newsletter".

Contributions are open to all researchers with personal experiences in the topics presented in the Newsletter. Papers are invited under the form of word processor files when generated by Commodore 64/128 software.

This is the last issue of Volume 01 and any subscription ends here. Subscriptions to Volume 02 are now open: please take note, readers!

"The Computer UFO Newsletter" is a special publication of the new Italian Center for UFO Studies, which foreign liaison office is:
c/o Dr. Edoardo Russo
C.so V.Emanuele 108
I-10121 Torino
ITALY



EDITORIAL

This is our last editorial.....for volume one of this Newsletter. The very first issue of CUFON came out about a year and a half ago, as an experiment: the interest shown by the international UFO movement has been quite good, even though only in recent times our presence has been remarked by some magazines.

CUFON circulation is limited to a number of sixty copies or so, but we think to increase our readership during the next months. But this could be a problem, as the Newsletter is xeroxed at times necessary to reproduce and compose by staples it is really enormous. We had had to renounce some personal interests in the UFO field for producing such a publication: furthermore, remember CUFON is entirely produced by a sole persona (this mad Editor !), with a very heavy employment of money and time.

All this led us to take a decision: CUFON will be published only in two issues in volume two, but each of them will be composed by sixty or seventy pages or so. Press quality will be improved by the use of special graphic programs, in conjunction with a normal word-processor (this is our present idea): where available, some illustrations will be presented. The aim is to reach a good qualitative level, as regards both contents and style. Obviously, the Newsletter will always need contributions from its readers under the form of articles, software and other material to be published or simply debated. Coming out once every six months people should have more time for prepare their own texts (we hope !). We know that matters developed inside CUFON are particularly specialistic ones, but each of you is interested in computers and has a certain knowledge about it: so the effort shouldn't be so hard !

Since now we open subscriptions to CUFON Vol. 02, we only way to finance the Newsletter (together with the offer of software and related print-outs). The rates are the following:

18,000 Italian lire surface mail and 24,000 for air mail, both to be paid ONLY by an International Postal Money Order (no check will be accepted, due to the very high cost of exchange, 8,000 lire). Please renew your own subscription and invite your friends/correspondents to get a new one. CUFON needs your contributions. Furthermore, we would like to know new ufologists having a computer, also in consideration of future common works to develop together: readers should provide this Editor with names and addresses of people they know to have a computer.

Important Notice

This brief note is essentially directed to the European readers. It is probable that within January 1987 we'll be finally able to establish an experimental UFO Bulletin Board System. First services should be electronic mail and the possibility to receive some files of documents, as well as available UFO software. During the period of six months or more the new service (probably working during night hours, from Monday to Friday) will be in an experimental phase. It will be useful to get experience with such a new application (US friends please don't laugh !), preparing new files to be made available to users. The

hardware employed in this project will be a Commodore 128 personal computer with a 1541 disk drive, at least in its very first beginning: in a near future, we are planning to upgrade the system on a more powerful PC, for example a Commodore AMIGA or a MS-DOS machine with hard disk. Obviously, it will depend on available financial resources (a little problem !).

To connect the system (the first one in the continent: it should be the European version of the US Computer UFO Network) a modem will be necessary. Knowing that most European ufologists haven't such a device, we want to invite them to buy one. The reason of this is obvious: such an application of computer technology to our subject is extremely interesting and it could be of great importance, if structured and managed in a suitable way: don't lose the opportunity of exchanging information in real time: now, it seems only a curiosity, but we are sure of its usefulness in a near future.

We don't want to speak about the BBS now: we only want to invite to purchase a modem, to be able to connect the system when it will be operative. Suggestions, opinions, critics and personal ideas about the possible use and applications of a UFO Bulletin Board System, as well as contributions in software (we are always looking for a really good BBS program), will be extremely welcome. When you will have the modem please contact us at once for more information. We hope to hear from you through the electronic mail of the system in the next months !

NEW Call for papers

As previously published in CUFON experimental issue, we present another call for papers. All readers using a computer for UFO activities are invited to prepare a text about their work, with a discussion about hardware, software and aims of the project. Graphics, tables and other topics related to the use and applications of computer in ufology are requested for an eventual publication, as well from people without any current computer project (their own ideas and suggestions could be valuable).

Papers have to be submitted in English language, typewritten on A4 sheets: eventual figures should be placed on a different page. Contributors having Commodore 64 or 128 computers should send their texts as a word-processor file, from "Easy Script" or "Superscript", if possible.

Remember :

Your subscription ends together with this issue. Renew it !

INTRODUCTION TO THE FORTEAN DATABASE

Bob Rickard

I am not an 'expert' on computers or systems analysis, nor, I suspect, are many of my colleagues in the study of the UFO or any other unusual phenomenon. When we decide to apply the personal computer (PC) to our study area we realise the problems facing us are truly formidable. Trying to make sense of the bewildering variety of computers and programs in terms of my own needs has been an education indeed. On this frontier one feels extremely isolated, vulnerable and even hesitant, especially when an impulsive purchase of computer or program can be very costly in wasted keyboard time as well as money. Consequently I applaud the establishment of this COMPUTER UFO NEWSLETTER which will go far in reducing the sense of isolation and vulnerability by sharing problems, experience, and skills. I am pleased to have this opportunity to explain the philosophy and design of TOAD, the fortean database operated by Archives for Fortean Research (AFR).

BACKGROUND

I first became interested in applying the PC to my Fortean studies in 1978. It was an exciting year which saw the beginning of the 'home computer' boom. Suddenly everybody was talking about the Apple, and several magazines appeared devoted to the new computers and programs arriving from the USA.

The temptation to buy was great indeed, but my hesitancy mounted as each month brought news of "better" computers and programs. Hardware and software seemed in a frenzy of development. Strangely, as prices tumbled their power and flexibility increased. (Eg. in 1984 a 5 megabyte external hard disk cost me nearly £ 2000 - today I can plug in a 20 mb hard disk, complete with controller, for under £ 500!). Also at that time there was a battle between a number of disk operating systems (DOS) - none of the manufacturing companies were interested in portability of programmes between machines, or machine compatibility, or user-friendliness. In those days database programs were non-existent - I would have to write my own application, which meant taking the time to learn Basic. Clearly the best policy was to wait.

The soundest piece of advice I was given, and which I always pass on is to ignore the hardware and work out your software need first. The first priority is to understand your data and what you want to do with it thoroughly - then search for the most applicable software - and only then look at hardware. You'll find that, given the constraints on your purse, finding the most suitable computer is relatively easy. So while I waited for prices to fall within my limited reach I decided to examine the nature of my data and processing needs.

In the course of publishing FORTEAN TIMES, and my own fortean research, I have accumulated tens of thousands of newspaper clippings, xeroxes, magazines and books, as well as research

notes, references and bibliographies. And these are being added to daily. Our collection has already gone past the point where we could find items easily using our own memory. Any archivist knows the problem of shelves groaning under the weight and numbers of items, and the organisational difficulties of proper indexing.

At first it was thought the advent of cheap powerful PCs would allow us to reduce the physical size and indexing of a library in one go. This is not the case: storage memory in magnetic media is still too expensive, even if we had the time to type in ALL the available data verbatim, indexing keywords as we go. This also applies to the storage of images on video.

For facsimile storage optical media are superior. The conventional approach is to use microfilm or microfiche and incorporate the frame reference into the computer database. There are already available programs allowing PCs to 'drive' a microfilm reader/printer. The latest research in optical storage attempts to develop the laser/digital technology of the hi-fi 'compact disc' (CD).

Until recently compact disks were 'read only' (CD-ROM), but major companies have announced breakthrough on systems which can read and write files to CD. As a sign of things to come, an interactive CD (CD-I) device called the 'Intelligent Archive' costs £ 3000 can store 115 mb per side, and for £ 17,000 or more 'gigadisks' are available offering 550 megabytes. This year will see a great battle between contending groups to establish a file-handling standard for CD-I, and only when that is settled can we expect a rapid market in cheap and portable optical memory.

Even when facsimile storage becomes affordable, there is the formidable problem of indexing. For our purposes it is not enough to know the address of an image of a clipping. Before each item is stored we would need to index the essential data, and the most logical place for this would be on a computer. So with or without associated facsimile storage (and we will stay with old-fashioned filing of paper for a while) there is a need for a conventional computer database.

FORTEAN DATABASES

I use the term 'fortean' as an adjective, meaning a thing, experience or phenomenon out of the ordinary (para-normal), which challenges conventional or consensus modes of rationalisation (eg. orthodox theory, or belief-system). In this view the UFO and ufology are categories of fortean interest.

The use of computers for what we might generally call fortean subject-matter came to the fore in the late 1960s. At that time the PC was a science-fiction dream and the pioneers in our field had to pirate the time and memory of mainframe computers owned by the companies or universities they worked for or studied at. As far as I can tell the earliest to try a computer correlation was Damon Knight, biographer of our eponymous mentor Charles Fort (1). In the early 1960s, when only big companies could afford the then room-sized computers, Knight convinced a friend at Bell Telephone Labs to key-in 1200 records derived from Fort's books. From this sample Knight claimed to have discovered a correlation of the frequency of anomalies with

the epochs of Mar's orbit.

After that the main applications of computers were to ufology, specifically - eg. the German MUFON's 'CODAP' system; Dr David Saunders' 'UFOCAT', begun in 1969 and donated to the Center for UFO Studies in 1975; and Ted Bloecher's 'HUMCAT', again for CUFOS in 1975. All of these systems were similar and borrowed from old punchcard-type methodology. They had rigid file design, with much use of codes or 'flags'; they could not cope with text entries well, and allowed limited cross-referencing and sorting.

A measure of the need for and application of UFOCAT can be had from chapter 20 of Allan Hendry's THE UFO HANDBOOK (1979), in which he says that UFOCAT contains "60,000 individual UFO entries". Just one year later, a CUFOS publication claimed UFOCAT had an estimated 100,000 entries (2) - quite a jump! If data can accumulate in one subject at that rate, my idea of a massive central database was folly. Even with a small number of dedicated fellow-enthusiasts, and available, affordable memory, the work could go beyond several lifetimes and still be incomplete. We have to set our sights on the achievable within our limitations - with foresight and skill we should also be able to lay the foundations upon which the systems of our successors can build.

TOAD - DESIGN CRITERIA

Our criteria are simple, arising out of bibliographical research and data usage. Besides the usual editing facilities the database should:

- 1) be capable of near infinite expansion at the record level.
 - 2) allow non-sequential entering of records and subjects.
 - 3) allow more than two files open so that complex reports can be generated from cross-referenced multiple files.
 - 4) allow for different types of data fields (numerical, character, logical and special date fields).
 - 5) allow alterations to file design while operational.
 - 6) allow browsing (ie. presentation of data in table format) and printing of screen contents, as well as the usual listing and print facilities.
 - 7) allow as many keywords as possible, for indexing.
 - 8) allow hierarchical sorting or indexing (eg. chronologically under alphabetical subject).
 - 9) allow the 'carrying' of field contents forward to a new record during data entry in cases of repetitive data.
 - 10) locate specific records swiftly.
 - 11) output files to a word processor.
 - 12) offer a range of 'error-checking' facilities.
 - 13) have a 'screen painting' facility for customised data entry and report forms.
 - 14) statistical and graphics facilities if needed.
- Etc. etc.

WHY IBM? WHY DBASE?

I looked at many systems and, truly, none of them were perfect. Given the amount of data I wanted to store, the type of

operations I wanted to perform, and the need for adapting files and fields to our specific needs, it was clear that proliferating gaming computers (Spectrums and Ataris, etc) and their simple software would be quite inadequate to the task - as I suspect many researchers are beginning to find out. (Today, companies like Tandy, Atari, Apple, Commodore and many others, who begun in the market with games computers or old 8-bit machines, have been forced to manufacture 16-bit MS-DOS machines to capitalise on the greatest variety of software written for this 'IBM-compatible' category.)

There is no ready-made system for a fortean database. The available software products were being edged out by a new approach - the database language. This allows the user to customise a basic system to his exact needs by a range of menu-driven or command-driven programming options. This type of "applications builder" has made it possible for idiots like myself, who have trouble learning Basic, to write quite sophisticated and powerful programs. Also modern database languages, like dBASE, have broken away from the early rigid hierarchical structures allowing data to be entered randomly and several files to be processed in complex relationships ('relational databases').

At the time the IBM-PC was clearly becoming the standard machine for which most software was being published, which was one of the factors which convinced me, contrary to my own advice, to go IBM before I decided on software. Obviously any serious contender would be available to IBM or compatible machines first. Also, being a lone, vulnerable end-user it was essential to have a machine which was as reliable as could be, and with good technical support if needed. There were fancier machines coming on the market every month, but I decided to play it safe. I paid about £ 1600 for my 128K IBM PC, with one disk-drive, two years ago, and I have added to it in that time. I recently upgraded to a 20mb hard-disk for £ 495 (+ VAT). If you shop around you can find IBM look-alikes, assembled from IBM-quality components (made in Hongkong or Taiwan), and a 20mb configuration similar to mine might cost under £ 1000. Good computing power is now within the reach of the serious researcher.

Having settled upon IBM, I realised I could apply the same principles to database selection. The choice had narrowed down to dBASE II and RBASE 4000, and a few others. Where RBASE offered a powerful range of facilities, it was still relatively unsupported in this country. dBASE was extremely well supported, but was severely limited by its inability to open more than two files at a time. Suddenly my wait paid off - the long-awaited release of dBASE III (dB3) allowed relationships between ten open files at a time, and an exciting range of functions for manipulating data. That tipped the balance, and I committed myself to dB3. There are indeed database systems now available which surpass dBASE in one or two functions, but few can match the power and range of its functions and commands. On such a far-reaching decision I thought it was better to play it safe and go for RELIABILITY. After all, I wanted a system which would be functional for at least 5 years (perhaps more). By then database systems will have developed to even greater power and sophistication - and because dBASE can

output standard ASCII files, its data can be transferred easily to any better system.

TOAD - THE PHILOSOPHY

Over the development period many systems and sources have influenced TOAD in my search for suitable systems of categorising and operating. I have emulated the good points of some and steered clear of the obvious pitfalls of others. It would be tedious, indeed impossible, to list them all, but I must say a sincere thank you to Anders Liljegren for his UFOCODE (3), devised for AFU; to Paul Jackson for his TASCAT (4), devised for the Tasmanian UFO Investigation Centre; to John Prytz for his painstaking efforts in devising a library subject heading and indexing system for the Australian Centre for UFO Studies (5); to Dave Fideler who sent a copy of the file-design used by Persinger for his 'space-time transients' analyses (6); to Bill Corliss for his monumental work both compiling AND classifying scientifically anomalous data (7); to the late Gray Barker whose UFO guide to FATE magazine showed the usefulness of a computer in a bibliographical application (8); to George Eberhart who seems indefatigable in his attempts to catalogue American Fortean; and to the early pioneers of UFO catalogues, Ted Bloecher and David Saunders, whose work is at the heart of the CUFOS database. You will notice that most of these efforts are in the field of ufology - without doubt it helps to work within a more easily defined field.

The first task was to find the basic units into which an incident can be broken down; the greater the number of basic units, the greater the number of connections can be made, and it is hoped that one day some of the permutations of these connections (or cross-references) will lead to real insights. I use two kinds of data groups: primary and secondary. PRIMARY DATA (subject, date, location, and reference) are shared by ALL records/entries. SECONDARY DATA, which is additional to the primary details, can vary considerably between records/entries. Secondary data fields tend to reflect the researcher's interests and would concern, say, witness-related data, the involvement of animals, meteorological conditions or anything else the researcher feels is important. (11) Each of the primary data field areas has its own problems and I'll deal with some of these under their headings below.

So it is a cardinal rule of database design to include as much flexibility as possible. A large number of small files can be processed more efficiently than a single massive file. For example: a record might only consist of primary data (subject, place, date, source) (10) and so there would be no need to tie up valuable memory in entering BLANK fields in any other subdivision. Therefore the structure of TOAD is a cluster of small specific files from which the data can be reassembled in a variety of combinations according to the inquiry.

Another sound principle of database design is to prohibit direct access to the database itself during data entry - to minimise the risk of corrupted records. Another problem is to reduce the number of BLANK fields. We have solved both these

problems by the use of an intermediary, temporary database. Imagine a file card with all the available fields or questions on it - it would look very much like a UFO sighting report form. Not all the questions are relevant so there is a lot of blank space, and if your database is set up likewise with one long record form there would be a lot of wasted memory. However, we can afford to design our temporary database like this only because it is temporary, and we need to be able to select from all the available fields for each new record. These long records with many blank fields can be checked and then processed as a batch - a filing program examines each one and 'posts' (or copies) each field to the relevant TOAD data files (see Fig.1), except where the field is blank, thus minimising memory wastage. And because the filing is automated the risk of corrupting the database is low (assuming no bugs!). After filing the temporary database is erased. The next time data entry is needed the program creates the temporary database, appends a blank form and assigns it the next available record number. And so on.

----- Fig 1 -----

```

*****
* Main data      *      * Main data      *
* entry form    * ----> * temp d'base * ---!      *****
*****
*****
* Source data   * ----> * Source data * ---!      *****
* entry form    *      * temp d'base *      *****
*****
*****
* Filing prog   *
* into TOAD    *
* see Fig.3    *
*****

```

(The basic flow of data entry into temporary databases, then filing into TOAD)

, The all-important link between each TOAD data file has to be unique. Originally I considered using a combination of date and subject code, but this would become confused if several incidents of the same kind happened on the same day. We have decided to use the automated filing program itself to generate the "next available record number" which then becomes the "unique key" through which all the relationships to that record are set.

SUBJECT

One big problem was how to deal with subject headings. The library-type of organisation (as proposed by Prytz, Eberhard and Liljegren) proves to be of little use in dealing with the complexity of cross-references needed in our database. When it comes to classifying such inter-related topics, as are typical for-ean phenomena, we have to build up a descriptive phenomenology from scratch. The for-ean arena is difficult to define: it is not a single subject, and many areas of our interest seem to be subjects which are 'interdisciplinary' and whose novelty lies in their "merging away into other subjects" as

Fort put it. Indeed, it was Fort himself who noticed that frequently it was not "things" but the "connections between things" which provided the challenging mystery.

Philosophically, Fort distrusted categorization because of the human tendency to use ready-made systems as substitutes for creative thought. Nevertheless, he saw the necessity of some kind of classification system with which to file and retrieve his notes and clippings. In 1919 he said he had covered his range of interests with about 1300 main headings. These were broadly descriptive in an unusual way: 'Harmony', 'Equilibrium', 'Supply and Demand', 'Metabolism' etc. I rather suspect that Fort revised and expanded his system, when, in about 1920 he burned his "40,000" notes, "because they were not what I wanted", and began collecting all over again. In any case, he left no description of either system. I have opted for more conventional headings because these will be more familiar to the user and any variations can catered for under the main heading and by cross-referencing. Besides, Fort's great concern for the "relations between things" is not lost, but enhanced, by the speed of the modern computer, and interest in conjunctions of data will be catered for in the query programme.

It is impossible to put a figure to the total number of subjects which will be available as data headings in TOAD. In compiling the subject catalogue for TOAD (9) we have listed several thousand legitimate topics with their assigned codes. As the subject catalogue comes into use I expect it will be expanded and AFR plan to release regular updates on new subject and their codes.

By 'subject' I mean a phenomenon as distinct from a 'case'. A case may be broken down into a collection of phenomenal incidents, and I take these 'subjects' to be one of the primary data units of my database. There is provision in the subject file itself for entering a 'case name' (which could be a witness name, a place name, or some descriptive phrase) where this is in common use, and a search on this field should allow the full reconstruction of the collection of incidents which go to make up a 'case'.

Subject codes are an integral part of the program design, because codes are faster and less ambiguous for the machine to use than relatively clumsy words and phrases we normally use. It is also easier to type in a five-digit code than a 60-character subject description. However, codes are not very 'user friendly', and there would be a need to substitute the text 'strings' when outputting to screen or printer.

A code has distinct advantages in the database proper. It needs minimal storage memory. When reports are made to the screen or printer, the subroutine looks up the code in a 'look-up' file (some people call this function a dictionary or thesaurus) and the appropriate subject title text is displayed or printed. This technique has another advantage: if changes or additions need to be made to a database in regard to 'subject', one simply changes or adds to the text in the look-up file. The actual database data is not touched. Also, by simply translating the text records in the look-up file a printout in a different language is possible.

One important facility must be the ability to conduct

hierarchical searches. EG. the subject catalogue has a general entry 'FALLS' for objects etc falling from the sky. This heading has 11 sub-headings (ie. falls of animals, artifacts, ice, inorganic material, liquids, oddities, conventional (eg. people who survive falling great distances), organic material, plants & plant material, stones, and slow falls). Each one of these has sub-sub-headings and so on. (see Fig.2) We can be specific to any

```
----- fig 2 -----
FALLS (271)
- Falls of Animals (272)
  - Falls of Amphibians (866)
    - frogs & toads (867)
    - tadpoles (869)
    - spawn (870)
  - newts & salamanders (868)
- Falls of Birds (871)
-
- Falls of Fish (874)
-
```

(Part of SUBJECT CATALOGUE, showing hierarchy and codes)

level we like (eg. falls of frogs = 867) or by coming up the hierarchy we can catch all the amphibian falls by searching for (866). Because each subject code will be stored along with the next highest number in the hierarchy (eg. to the database the full code for falls of birds is 871-272, ie. a sub-heading of Falls of Animals) in the same 'look-up' file as the subject heading text string, we can cater for searching on specific headings OR general headings.

DATE

I wanted more than a simple 8-character date field. dBase has a special data field for dates which can be manipulated by special date functions. For example:

1) dBase has functions for returning the day of the week in word form, and this can be automatically inserted into a field in the record upon input of a numerical incident date.

2) the AGE of a witness can be automatically calculated upon input of date of birth; and vice versa, a DATE OF BIRTH from input of age. dBase offers a number of date manipulation functions, like the ability to extract numerical values from a word date so that another date can be computed (eg. if you have a date in the form <29/May/1986> and wish to add nine days.) In many reports we are given two datums (eg. the incident date and the age of the witness), so it is easy to calculate the third datum (eg. date of birth).

dBase III appears to be configured to use only dates after 1900, and this was a serious disadvantage when dealing with pre-1900 data. However one young whizkid showed me a routine which by-passed this limitation in the old dBase. The problem does not arise in the new dBase III+, which includes a new command (SET CENTURY) which toggles between '20th century only' (last two

The problem of 'vague dates' is another headache. I have decided to make provision for two alternatives in the data entry stage, when the operator must choose between a full or a vague date entry screen. The 'full date' is one where all the numerical components of the date are known, and so it is on this part of the program that such date calculations as those mentioned above can be made. The 'vague date' caters for such data as is known, including seasonal dates: "the first Tuesday of March (year unknown)", "sometime in summer 1945", "in the early 1870s", etc. We do this by the use of fields for both codes and text.

```

*****
* Subject *
* (thesaurus) *
*****

^
|
|
|
|-----
* X-Ref *
*****

^
|
|
|
|-----
* X-Ref *
*****

^
|
|
|
|-----
* X-Ref *
*****

*****
* Subject *
*****

*****
* Place *
*****

*****
* Date/Time *
*****

*****
* Citation *
*****

*****
* Meteorology *
*****

*****
* Witness/etc *
*****

*****
* Animals *
*****

-----
* Others *
-----

```

Arrows indicate the direction of a relation. NB: 'Others' means that new data files may be added to the structure.

References and sources provide the most difficult of all the problems. Because one of the researcher's tasks is to generate a bibliography, we must have some general link between subject and reference, as well as the standard link between incident record and reference. EG. we might want a list of our source material on a general topic, say 'psychological theories of UFOs', rather than an incident-related reference. It is possible, but cumbersome, to search the whole database for

records with the appropriate subject code, then to compile a list of source references from those records. It is far easier, at the time of the automated filing, to enter the subject code and the source reference into its own file, and the search we have described would begin there.

Potentially more complicated is the great variety in types of sources (books, news-clippings, films, photographs, investigation reports, etc). To begin with we must distinguish once more between primary and secondary data. The relationship between primary and secondary source files is similar to that between primary and secondary data groups explained above. PRIMARY SOURCE fields concern self-contained sources, such as books, in which the contents have a fixed relationship to the title. In SECONDARY SOURCES this relationship is variable: eg. the title of a newspaper, journal or annual proceedings stays the same, but the contents vary with each different issue date. Therefore secondary source records must include a primary source code.

We must also cope with two different type of multiple referencing. One incident or topic may have a great many reference sources - conversely one source (eg. Fort's COMPLETE BOOKS) may refer to a great many incident or topics. To avoid the massive repetition of reference details (a waste of valuable memory) we apply the trusty technique finding the smallest discrete data groups. Thus we have a cluster of files dealing with primary source, secondary source, and a cross-reference file (x-ref) which links the bibliographic files with the incident data files. Figure 3 shows how the various TOAD files relate to each other.

ONE-TO-MANY / MANY-TO-ONE

The structure described gives us the flexibility to make the following relationships:

- * a single subject code may be linked to many incident records (allows the reconstruction of 'cases' or compilation of lists).

- * a single subject code may be linked to many primary and secondary source records (for bibliography of a subject).

- * a single incident may be linked to many primary and secondary source records (for bibliography of an incident or 'case').

- * a single date may be linked to many incident records (for compiling a chronology or list of events for one day).

- * a single place record may be linked to many incident records (for compiling a gazetteer of phenomena).

- * a single primary source record (eg. a newspaper) may be linked to many secondary source records (articles or clippings), each of which may be linked to many incident records or subject codes (for checking all citations from one source).

- * and a great many other combinations - eg. a correlation could be sought between a witness name, a day of the week, a type of animal, a location and certain types of phenomena.

- * etc, etc.

PROGRESS

I have not said too much about the actual fields - these will be fully described in a future paper because they are currently in a development stage (Eg. I am trying various ways of storing and manipulating dates and I have the choice of putting one kind in the main data entry form and other kinds in the actual data file. So to present a listing here would be both premature and may be misleading.)

Progress has not been as fast as I would like. Once the careful choice of computer and database language, I have had to learn dBase (and then dBase+) to advanced programming level. Had I been able to sit down and devote myself to the problems alone the database would be operational now - but alas I have only been able to snatch an hour or two when I could. Writing programs is the easy part - debugging and proper testing takes up most of the development time. Here is the state of play:

- * Main data entry program - now in third version and undergoing a major re-write to improve presentation, ease of use, and execution. Near completion.

- * Source data entry program - Program functional, and source database is in operation.

- * Subject thesaurus program - program functional but nearly 1000 records damaged during a disk-crash! Reconstruction underway.

- * Filing program (which takes data entry forms and files them into the TOAD database proper) is near completion.

- * Maintenance programs (for updating database, or deleting duplicated records, etc) exists in pseudo-code only.

- * Full range of query and reporting programs are planned, a few of which exist in pseudo-code only. This is next phase of development.

USE OF TOAD

Although going "on-line" is a very attractive idea, it is at least 2-3 years away. Before it happens the inquiry procedures and data security problems must be improved.

In the meantime AFR hope that ANYONE can join in data entry, even if they do not have a computer. We plan to make a simple 'kit' available of data forms and a guide (including codes). AFR also plan to publish a subject code catalogue (see note 9), and a catalogue of sources. It is anticipated that these last two (the subject and the source catalogues) will need revision every year to accomodate the wealth of new material. Members of AFR will be offered them at reduced rates.

Beyond that, our plans are to purchase a dBase compiler so that cut-down versions of the database structure may be made available to researchers with micro-computers.

We have put a lot of thought and money into developing TOAD. We have emphasised flexibility. As TOAD is used more and more, its structure will undoubtedly undergo modification: field sizes may be altered, new fields and files added, and new query or reporting programs written. Because of its power and growth

potential we hope more and more researchers will be attracted.

Individuals or groups could keep copies of their own data on TOAD as backup, or use TOAD facilities for analysis. Some of these services may have to be paid for, with members getting concessionary rates: but nothing has yet been decided on HOW to administer AFR in this respect, although something on the lines of Hilary Evans' BOLIDE information network is attractive.

There is a real prospect of establishing common standards for fortean databases by developing TOAD in cooperation with other groups - and I would be happy to discuss this with anyone. UFOlogy is a major field and a complex one, but it is one of many which share similar data problems, which we believe TOAD addresses directly.

NOTES

- 1 - CHARLES FORT: PROPHET OF THE UNEXPLAINED by D. Knight (Doubleday 1970).
- 2 - PHYSICAL TRACES OF UFO SIGHTINGS (CUFOS, 1980). This printout from UFOCAT has "nearly 2200" entries.
- 3 - UFOCODE: A CLASSIFICATION SYSTEM FOR UFO REPORTS, UFO LITERATURE AND REFERENCES ON UFO-RELATED SUBJECTS: VERSION 1 - DEC 1983 (Archives For UFO Research, Box 11027, S-600 11 Norrkoping, Sweden). 'UFOCODE: A Classification System for UFO Research' by Anders Liljegren, AFU NEWSLETTER 26 (May-Dec 1983).
- 4 - TASCAT: TASMANIAN UFO COMPUTER CATALOGUE - 1985 (Tasmanian UFO Investigation Centre: Box 99, North Hobart, Tas 7002, Australia.)
- 5 - A SCHEME FOR THE CLASSIFICATION OF UFO AND RELATED INFORMATION by John Prytz (paper for UFOCON 7, Hobart, Tasmania; 1983; unpublished). SUBJECT HEADINGS AND INDEXING TERMS RELEVANT TO UFOLOGY AND RELATED DISCIPLINES by John Prytz (paper for UFOCON 8, Sydney, NSW; June 1984; unpublished). John's papers contain an astonishing amount of work and thought on classification problems.
- 6 -, SPACE-TIME TRANSIENTS by M.A.Persinger & G.F.Lafreniere (Nelson-Hall, 1977).
- 7 - A CATALOG OF GEOPHYSICAL ANOMALIES - A projected series of about 25 volumes, of which 5 have been published so far. Compiled by William R. Corliss (Sourcebook Project: Box 107, Glen Arm, MD 21057, USA.)
- 8 - A UFO GUIDE TO FATE MAGAZINE by Gray Barker (Saucerian Press: Box 2228, Clarksburg, WV 26301, USA; 1981.)
- 9 - A GENERAL CATALOGUE OF FORTEAN SUBJECTS FOR THE FORTEAN DATABASE (Preliminary edition), compiled by Bob Rickard (Archives for Fortean Research, 1 Shobury Rd, East Ham, London E6 2AQ; May 1985). This preliminary edition had only 12 copies, and was selectively circulated for criticism and improvement. The next edition (the 4th) will be the first public edition, and will contain subject codes for TOAD.
- 10 - Our editor, Maurizio Verga, has called these fields groups "fundamental data" COMPUTER UFO NEWSLETTER v1 n3 (Jan 1986) p4f. What I call 'secondary data' could be viewed as an expansion of his 'Classification', 'Evaluation' and 'Notes' fields.
- 11 - My initial selection of secondary fields is personal, of

course, reflecting my own research interests. I was not worried about overlooking potentially important data, because the facsimile of the item could always be referred to and the new information included at a later stage. As other researchers join the network, each with their own interests, I expect the range of fields to be expanded - and the ability to alter a database structure in this way, while the database is operational, is one reason I went for dBase.

Bob Rickard

(Bob Rickard is a well-known Fortean and UFO researcher, editor of the highly-respected magazine "Fortean Times". He is very interested in the use of electronic data processing in our subjects (especially Fortean), due to the huge mass of data he has to handle. T.O.A.D. seems a really interesting database system for such a kind of event and we are looking forward to receive further papers about its development and concrete applications. Comments and suggestions from readers are invited, obviously.

We remark that dBASE III is also the database system of SCANCAT, the Scandinavian Catalogue of UFO events managed by Anders Liljegren. The report file has more than 1700 recorded sightings at moment (early September 1986), most of them related to 1930's "ghost fliers" and 1946 "ghost rockets". The record structure includes 36 fields, with a total length of 185 characters: most information are inserted under the form of codes. Liljegren has a Victor VPC 15, a PC IBM/XT compatible machine and he is planning to write a paper for CUFON about his own work.)

|||||

PROBABLE MEETING AMONG CUFON READERS

Next year, a workshop exclusively devoted to the possible employments of computer in ufology should be held during the fourth international BUFORA UFO Congress (August 23-25), in London. It could be an excellent occasion to organize an informal meeting among all ufologists interested in computer applications, to debate our subjects and exchanging ideas and material. This Editor will attend that Congress and he hopes to meet most of you in that occasion. Moreover, there will be a similar workshop during the international ICUFOS Congress (May 1987, in Turin) where a large part of it will be devoted to computers and ufology: in view of foreign participants, a service of translators will be active. Further information on next issues of "The Computer UFO Newsletter", which will promote such meetings.

UFOCOMFILE

- Introduction

UFOCOMFILE seems to be an interesting project for the storage of a national casuistry on computer. The idea was good, even though we don't agree much with the great mass of coded data inserted in each record. Unfortunately, this work headed by Australian researcher Andrew Cole (who is the owner of the Alpha-Micro 100 where the system has been implemented at the moment) hasn't received a concrete contribution from local groups and investigators. This has practically caused the failure of the project, as regards what we know.

We think the cause of this is quite simple: the whole project (which aims were actually interesting) had an old conception of the use of E.D.P. It was founded on a sole big computer and contributors had to supply compiled cards (related to their own cases) to the operator of that machine. All this involves a lot of problems, such as the compilation of cards (a very boring operation !) and a scarce psychological participation of contributors to the project. Moreover, a work of such a kind needs a lot of time, also because there is only one man in inserting data in the computer. According to our own point of view, the most suitable solution for the computerization of a national casuistry is the establishment of a "network" of personal computers, owned by different ufologists. Each of them prepares a piece of the casuistry (for example, all sightings happened in its own province) employing a common database program. The resulting files can be collected together and then passed on a more powerful computer to carry out statistical and other kind of analysis.

This is only our personal solution about the storing of UFO sightings on computer, in view of a national catalogue. We are completely open to any other proposal and suggestion able to establish a positive debate. For the moment, here is a summary of UFOCOMFILE structure and activity, drawn from Andrew Cole's booklet "The Australian UFO Computer File - Instructions and codebook".ÆEDITORU

UFOCOMFILE

1. THE PURPOSE OF THE COMPUTER FILE

The computer file is designed to fulfil two main objectives:

(A) To act as a quick-access filing system for all Australian UFO reports. Access to be available to those groups who contribute data to the file, as well as other persons or groups (private researchers, etc..) at the discretion of ACUFOS. Private individuals or groups who are not members of ACUFOS. Private individuals or groups who are not members of ACUFOS will be charged an "operating and materials" fee for

supply of information from the computer.

(B) To provide the means of conducting research into the UFO phenomenon in areas where comparison of data and statistical evaluations are envisaged.

2. OPERATION OF THE COMPUTER

The Australian UFO Computer File was originally set up as a punched-card file for operation on a UNIVAC 1004, it was then migrated to disc at Automatic Data Service in Canterbury N.S.W. on a I.C.L. 1902A. This was the first time that an external cost was associated with the file.

The whole file architecture was then modified to allow for complex analysis and was moved to a disc based multi-user system at Australian-Alpha-Micro, the programmes being written using a very sophisticated form of compiled Basic.

Due to various personal commitments, the file ceased operation from 1981 to 1984 when houses were mortgaged and a fairly large Disc based, multi user, Alpha Micro system was purchased. This allowed for future networking possibilities. Also, ownership meant the end of borrowed machine time and programme re-writes for system migration. Electricity and paper costs are the only overheads.

3. DATA FILE FORMAT ORIGINS

The original file format was finalised in December 1976 having being drawn by myself and Michael Smyth. Input from Dr. David Saunders and Fred Merrit was obtained so as to establish some compatibility with Dr. Saunders' UFOCAT. Compromise was reached with exceptions and borderline cases of category selection being kept to a minimum without the whole system becoming an unworkable megalith.

It is important to note that previously entered reports that may have had un-codeable details may now have codes available for entry of this new information. In fact, any new information may be added to an individual record up to the new maximum of forty different parameters in total.

The original record format compromised of one or two eighty column punched cards, consequently, the record was set up to fit into this format. The ideal size of a record on a disc based system is an evenly divisible portion of a disc sector which, on the Alpha Micro, is five hundred and twelve records, so one hundred and twenty eight was chosen. This will also suit various floppy-disc based microprocessors. At this point in time the only method of transferring the file from the Alpha Micro is via a simple character oriented programme down an asynchronous link. All binary compacted numbers will have to be expanded prior to transmission and re-compressed at the other end. The record format as it stands comprises of forty four characters, set up as forty pairs, in the parameter table. The remaining four characters are used for monitor flags.

4. DATA RECORD FORMAT

This format has been set up in the basic belief that ALL data obtained from the investigation of UFO phenomena is of equal

importance, therefore no apparently useless information should be sifted out or deleted. For any serious analysis to be carried out, it is of outmost importance that all identified as well as unidentified reports are entered. Only if this is done will the calculation of Probability Percentages be possible to enable our extremely limited field investigation capabilities to be utilised more effectively. This has not been done in the past and a large backlog has developed which must be entered if any usefull and for the sceptics, believable, statistics are to be forthcoming.

5. GENERAL DATA SECTION

CHAR. 001-009 These form the FILE NUMBER and are split into three groups:

1-2 SOURCE GROUP CODE (from coding tables). Each group has its own individual code which is to be used for reports that have been investigated by that group. This may include reports from outside that state, but do not include reports taken only from newspaper accounts. News reports that have not been directly investigated by the group are coded according to the "Other Source" code for that state.

3-6 REPORT YEAR. This is the year that the report was compiled, not the year of the sighting, for example a report received in 1977 of a 1974 sighting.

7-9 REPORT NUMBER. This indicates the number of the report in that year, for example, the first report for 1985 will be 001. The numbering of reports is also irrespective of the chronological order of the sighting dates for that year.

CHAR. 010 SIGHTING IDENTIFICATION LEVEL. "Possible identified" refers to reports that appear to have a conventional explanation but for which there is some doubt: for example, an object displaying standard aircraft lighting, making aircraft sounds, etc..., but which cannot be positively identified due to some unusual feature. This category level also includes "ambiguous" objects that appear to have no explanation, but, because, of some reason such as insufficient information, cannot be definitely classed. The other two levels, "Positively Identified" and "Positively Unidentified" are self explanatory.

CHAR. 011 SIGHTING CATEGORY NUMBER. Hynek's original classifications have been expanded and modified slightly to adopt the following meanings:

- * "Nocturnal" refers to sightings other than CE's that occur at night.
- * "Daylight" refers to sightings other than CE's that occur in daylight.

Here we have grey areas, i.e. when does night become day and vice versa. Sunrise and sunset appear to be of little use here as the sky is bright after sunset and also before sunrise. The best solution would appear to be to regard it as night if it is substantially dark and stars would be visible. The decision between day and night is probably best left to the discretion of the investigator.

- * "Instrument readings or trace only" refers to cases such as trace cases without a directly associated sighting. Cases of anomolous photographic images also fall into this category, i.e. cases where a printed image was not seen when the photograph was taken.

- * "Close encounter type 1" and "Close encounter type 2" both remain as originally defined by Hynek.

- * "Close encounter type 3" has been replaced with four categories, as suggested by David Webb (see his "1973 - Year of the humanoids") and one by Mark Moravec, investigator from Sidney. These are:

"Entity Report, class A" (ERA) covers a report in which an entity is seen inside, entering, leaving or in close proximity to an object, implying "occupancy". This entity can be human, humanoid, anthropoid or monster-like in appearance. The distance between observer and entity is covered by parameter code "J". This covers CE3, occupant and Webb's A-C degree of association.

"Entity Report, class B" (ERB) covers a report where an entity is observed without an associated UFO, but where the entity is similar to a type which has been reported previously in association with a UFO. This covers the term "humanoid" and Webb's "D".

"Entity report, class C" (ERC) covers a report where an entity is observed, but no association between the entity and any UFO phenomena can be established at the present time. This covers Yetis, Yowies, Black Panthers, etc..... and Webb's "E" and anthropoids class.

"Entity Report, class D" (ERD) covers reports of seemingly telephatic, audio, invisible or visible beings communication. This includes all "contactees" and "bedroom invaders" such as the Stuart case of New Zealand, 1954.

"Entity Report, class E" (ERE) covers a case where a witness enters a UFO involuntarily and undergoes a medical examination or a similar technique, administered by the UFO entity/s. This includes all abduction cases.

Due to the limited number of "occupant" reports, it would be of little use to attempt to include description breakdowns of entities. This will probably be the subject of special investigation in the future.

CHAR. 012-015 YEAR OF SIGHTING. This is the year of the actual sighting and not the year of the report.

CHAR. 016-017 MONTH OF SIGHTING. Self explanatory.

CHAR. 018-019 DAY OF SIGHTING. Self explanatory.

CHAR. 020-023 LOCAL TIME OF SIGHTING. Care should be taken in States where Daylight Saving is used. Daylight Saving Time must be converted to Local Standard Time and this figure used. Note that sightings occurring between 0000 hrs and 0100 hrs during daylight saving will in fact be somewhere between 2300 hrs and 0000 hrs the previous day in local standard time. Convention decrees that the sighting time be given as the time of sighting commencement.

CHAR. 024-027 SIGHTING TIME AS G.M.T. G.M.T. is used as a standard time reference for comparison data, so the correct conversion must be applied to the Local Standard Time of the sighting and this figure used here. Note that here, also, the date may change but this will be calculated by the computer if it is required for analysis.

CHAR. 028-037 PLACENAME OF SIGHTING. This field is used as a visual cue when scanning through printouts. Three different people would abbreviate the same place in three different ways, so it is not used for comparison data. The name of the nearest landmark, creek, town or maybe even homestead could be inserted in this ten character field.

CHAR. 038 STATE OF SIGHTING. This character is used for the code of the State in which the sighting occurred.

CHAR. 039 QUANTITY OF WITNESSES. "Instrument readings only" refers to cases (other than false alarms) where some form of detection equipment records an anomaly but there were no observers or sighting, as in the case of automatic recording devices. "Trace marks only" refers to ground marks, broken tress, indentations, etc.... that have no apparent sighting association. "Photographic evidence only" refers to anomalous photographs or photographs from unattended cameras.

CHAR. 040 AUSTRALIAN MAP DIVISION. This forms the first part of the Map reference system. Saunders uses latitude and longitude but it was felt that a "grid" has certain advantages. Lat. and Long. infer a high degree of accuracy in locating the position of an object and in some cases places a false degree of accuracy to a sighting data (no pun intended !). Using a grid system avoids this problem by assigning an area in which the object was seen. However, there will always be borderline cases, but it must be remembered that we are not trying to find a substitute for the data in a report, only a means of expressing it in a form that can be manipulated by a machine. For instance, flap predictions only indicate an area of high probability rather than a precise point.

CHAR. 041-042 MAJOR GRID REFERENCE. Refer to map appendix alleged with the codebook.

CHAR. 043-044 MINOR GRID REFERENCE. Refer to map appendix alleged with the codebook,

CHAR. 045-124 PARAMETER TABLE. These characters are set out in groups of two, with each combination having an unique meaning. Due to the large amount of data in a written report, data compression is required and it is achieved in this way. The Codebook contains all these parameters which are added to from time to time.

CHAR. 125-128 MONITOR FLAGS. These are only used by the computer during certain programmes and also change their meaning depending on the programme using them.

+++++

- Offer of Software

On CUFON Volume 02 issue one there will be the new complete up-to-dated "Offer of software" section, with new interesting items for popular personal computers. Several print-outs of UFO catalogues will be made available at cost price.

New Subscription

This is the last issue of Volume 01 and your own subscription is finished. Volume 02 will have two large issues published five/six months one from the other. Contents will be improved more and more together with press quality (inside the limits of our budget !). Subscriptions to Volume 02 are open at the following rates:

- 18,000 Italian lire via surface mail
- 24,000 Italian lire via air mail

Payments must be made only by International Postal Money Order payable to "Maurizio Verga". No check will be accepted, due to the very high exchange cost (8,000 Italian lire).

Renew your own subscription ! CUFON needs your support.

- Andersen, Per
Project UFODATA
Volume 01 Issue 03 Pag. 06-15
- Bianchi, S. & Di Ruzza, L.
A program for the investigator
Volume 01 Issue 03 Pag. 16-18
- Breysse, Denys
Checkpoint number two
Volume 01 Issue 04 Pag. 03-05
- Breysse, Denys
Becassine: checkpoint number three
Volume 01 Issue 05 Pag. 08-11
- Breysse, Denys
A database for entity cases
Volume 01 Issue 03 Pag. 19-21
- Cabassi, Renzo
The structure of a database for UFO events
Volume 01 Issue 04 Pag. 10-14
- Cole, Andrew & Ed.
UFOCOMFILE
Volume 01 Issue 06 Pag. 16-20
- D'Orazio, P.G. & Lollino, G.
Suggestions for a methodology in computerization of cases and reports
Volume 01 Issue 02 Pag. 09-13
- D'Orazio, P.G. & Lollino, G.
Suggestions for a methodology in computerization of cases and reports
Volume 01 Issue 01 Pag. 07-09
- Ed.
EDITORIAL
Volume 01 Issue 01 Pag. 03-06
- Ed.
EDITORIAL
Volume 01 Issue 02 Pag. 03-04
- Ed.
EDITORIAL
Volume 01 Issue 03 Pag. 03-05
- Ed.
EDITORIAL
Volume 01 Issue 04 Pag. 01-02
- Ed.
EDITORIAL
Volume 01 Issue 05 Pag. 01-02
- Ed.
EDITORIAL
Volume 01 Issue 06 Pag. 01-02
- Ed.
Computer & Ufologists: a preliminary census
Volume 01 Issue 01 Pag. 17-18
- Ed.
Offer of software
Volume 01 Issue 01 Pag. 14-16
- Ed.
A preliminary bibliography about applications of computer in ufology
Volume 01 Issue 04 Pag. 06-08

- Ed.
Census of computer & ufologists: an up-to-date
Volume 01 Issue 04 Pag. 02
- Ed.
Towards a common database for UFO events
Volume 01 Issue 02 Pag. 20
- Ed.
Computer & Ufologists: an up-to-dated census
Volume 01 Issue 02 Pag. 15-17
- Ed.
Offer of software
Volume 01 Issue 02 Pag. 14-15
- Ed.
Census of ufologists' computers: anew up-to-date
Volume 01 Issue 05 Pag. 20
- Ed.
Offer of Hardware
Volume 01 Issue 05 Pag. 11
- Ed.
Computer UFO Network
Volume 01 Issue 05 Pag. 03-04
- Ed.
Rete Ufologica Computerizzata
Volume 01 Issue 05 Pag. 02-03
- Jackson, Paul
Tasmanian UFO catalogue programme notes
Volume 01 Issue 02 Pag. 18-20
- Kampherbeek, H.
Horoscopes in ufology ?
Volume 01 Issue 05 Pag. 05-07
- Kampherbeek, Henry
UFODOC: a system for Dutch/Belgian cases
Volume 01 Issue 02 Pag. 05-08
- Rickard, Bob
Introduction to the Fortean database
Volume 01 Issue 06 Pag. 03-15
- Toselli, Paolo
Project of annotated bibliography
Volume 01 Issue 04 Pag. 09-10
- Verga, Maurizio
Sky Travel - How to identify astronomical IFOs
Volume 01 Issue 04 Pag. 15-16
- Verga, Maurizio
A computerized catalogue of Italian close encounters
Volume 01 Issue 01 Pag. 10-13
- Verga, Maurizio
Using Superbase
Volume 01 issue 05 Pag. 19-20
- Wootten, Mike
UFO data processing
Volume 01 Issue 05 Pag. 12-18